Leakage Current Analysis

Hao Chen, Latriese Jackson, and Benjamin Choo
ECE632 – Fall 2007
University of Virginia
<hc3r>, <ldj8e>, <byc6j> @virginia.edu

Abstract

Several common leakage current reduction methods such as power gating, dual threshold voltage, and body biasing were applied to examples models of carry look-ahead adder, carry bypass adder, and 6-bit decoders. The simulation results show that power gating and dual threshold voltage may be an effective method of leakage reduction while body biasing should be used as a supplementary method for leakage reduction.

Key Words

Leakage Current, CMOS scaling, power gating, dual threshold, body biasing.

I. Introduction

As technology scaling in CMOS technology advances, the portion of leakage current compared to the total consumed power of a CMOS based circuit is becoming larger and larger. Thus, many methods have been proposed to reduce the leakage current. However, as new CMOS scaling technologies arise the effectiveness of these methods has changed. In this paper, we examine several key leakage current reduction methods. First, we looked at power gating which uses a header and a footer transistor to create a virtual V_{DD} and GND to cut off the leakage path. Second, we looked at dynamic body biasing which created a bias in the body of both PMOS and NMOS transistors to reduce leakage. And finally, we looked at the dual V_t method which used transistors with different V_t for different delay paths.

We examined the advantages and disadvantages of each method, and simulated these methods in various application examples. By doing so, we were able to compare the effectiveness of each method.

II. Leakage Reduction Methods

1. Power gating

Power gating is one kind of dual threshold voltage technology that is greatly used to reduce the standby mode sub-threshold leakage. The scheme and operation of power gating is illustrated in Fig 1.

The logic block is connected to the sleep transistor considered as virtual power supply.

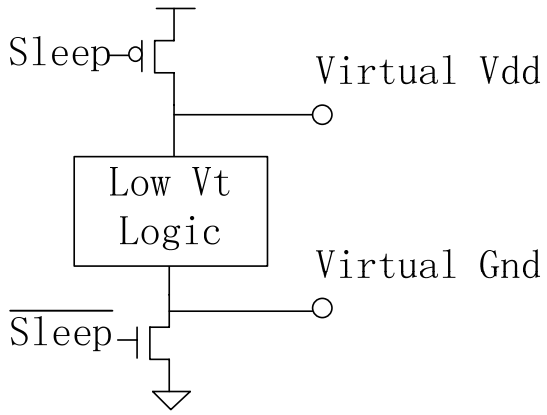


Fig. 1. MTCMOS circuit structure. [1]

When the block is in the standby mode, the sleep signal will shut off the actual power supply to the logic block. Thus the leakage current of the function is limited by the sleep transistor. When the block is in the active mode, the sleep signal turns on the sleep transistor, the logic block then is connected to the virtual power supply. In Fig 1., both the header and the footer transistor is provided, actually only one of them is needed in a block, if the circuit is completed complementary. The implementation is easy for the sleep transistor by simply adding a high V_t header or footer to a logic block.

In the active mode, the sleep transistor can be model as a serial resistor and a serial capacitor when sized properly.[1] Current flowing through the parasitic resistor will induce a voltage drop, which will decrease the gate-drive voltage and increase of V_t which is cause by the body effect. Both these effect will increase the transient time. So using power gating will degrade the circuit performance. A larger device will have a smaller voltage drop V_x, which will have less delay. But a larger device will not cut the shut off the leakage current significantly in the standby mode. To optimize the circuit, people need to design the sleep transistor as large as possible in the tolerance of delay.[2]

Another drawback of power gating is sometimes it may have a reverse conduction phenomenon.[1] When a NMOS in the logic is conducted and discharged before some other NMOSs begin to discharge, the current may flow from the other NMOS to the previous NMOS and pin the output to V_x. This will reduce the noise margin.

2. Dynamic Body Biasing.

The scheme of body biasing is illustrated in Fig 2.

Dynamic body bias is another effective technology to reduce leakage current. In the standby mode, the P body line is larger than V_{DD}, and the N body line is smaller than GND. Both NMOS and PMOS have a reverse body bias (RBB) between the source and body, and this will increase the threshold voltage thus reduce the leakage. In the active mode, the P – body line is smaller than V_{DD} and N- body line is larger than GND.[2] Both NMOS and PMOS have a forward body bias, and this will reduce the threshold voltage, thus have a large active current to shorten the delay.

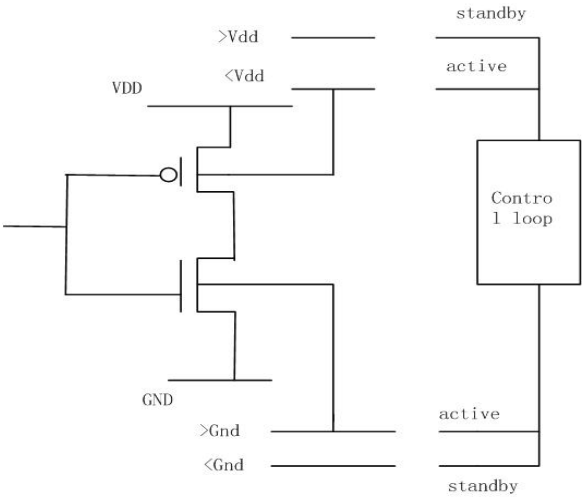


Fig 2. Dynamic Body bias

The drawback of this method is that it needs a voltage control loop, so costs extra area and extra energy in charging the substrate or well when enter and leave the standby mode. Because the original source body junction is a reversed PN junction, which has a large capacitance. In the modern COMS technology, the Halo structure is widely used, which has a highly doped body region around the source region. This makes the junction voltage drop is very steep. If people give a very large RBB, the band to band tunneling leakage current will exponentially increase. So the dynamic body bias control is very limited.

3. Dual $-V_t$ Domino Logic

In the power gating technology, a serial high V_t header or footer is placed to just cut off the leakage in standby mode. And it needs appropriately sizing the sleep transistor to balance the delay and the leakage. Usually it is not easy to size the transistor correctly. People also prompt anther dual-V_t circuit. The idea is to use both high V_t and low V_t transistors in the logic block. Use the low V_t transistor to design the critical path, and high V_t transistor to design the parts with non-critical paths. With the increase number of the critical path, this method becomes less efficient. And also people should carefully design the input pattern of the low V_t transistor, which make it really difficult to implement this method.

Dual V_t domino logic is illustrated in Fig 3.

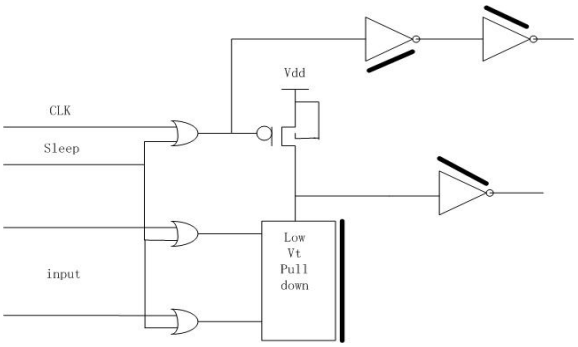


Fig.3 Dual V_t domino logic gate with Low-V_t device shaded.

Dual V_t domino logic can solve the drawbacks mentioned before. In the standby mode, the low V_t pull down network is turn on so internal node is discharge into ground, and PMOS High V_t transistor is the only source of leakage current, which is much smaller compare with low V_t device. In the evaluation mode, the current discharges through the low V_t pull down network. So the evaluation delay is equal to the low V_t circuit. In the pre-charge mode, the current charges through high V_t transistor, which have a large delay time.

III. Simulation result

1. 64-bit Adder

We built two static CMOS 64-bit carry bypass adder. The first adder channel length is 45nm and the second adder channel length is 32nm. This carry bypass adder was built by first designing a Full Adder then using that to design a four bit adder with a multiplexer (carry propagation). A symbol for the carry propagation was made and combined to create the full 64-bit adder. By designing and combining smaller portions of the adder to make the entire 64-bit adder, it was easy to test individual parts of the circuit to see how much leakage was create at individual parts in the circuit. In order to simulate power gating, 64- bit adders had to be created using low voltage transistors in the body and added high V_t device between the body and ground.

For the simulations, we started by simulating a single full adder. We simulated all the possible inputs to the full adder to see which produced the most leakage. Once we knew which inputs caused the most and the least leakage, we simulated our multiplexer independently to estimate the amount of leakage that it may produce if it was added to our circuit for vector forcing. Once the multiplexer was added to our circuit, we simulated the full adder with the multiplexer. Next we simulated the entire 64-bit adder without and with the vector forcing. The vector that the circuit was forced to when it was not in use was decided by the taking the vector caused the least leakage when a single full adder was tested.

Next, we applied power gating. As a result of the vector forcing testing, we had the simulation results for the normal full adder and 64-bit adder. The sizing of the high voltage transistor we decided to keep the same size as the transistors in the circuit seeing as that's the transistor sizing we wanted to test. We looked at different threshold values to see which values created that least leakage and did not inhibit the circuit.

We decided to see what would happen if we applied multiple methods based on the best results of the simulations to see what would happen.

The situation in which all the inputs into the full adder equals zero is the lowest leakage vector for the 32nm and 45nm full adder. Therefore, the vector the multiplexer used for vector forcing will force the full adder to zero when it is not in use. In the 45nm case the circuit did not seem to benefit from vector forcing because the leakage caused by the additional multiplexer outweighed any benefit it could have been to the circuit. The simulation results are summarized in Table 1.

Table 1. Leakage current comparison for 64 bit adder

Condition	Leakage
45nm Normal	250 uA
45nm LVT	11nA
32nm Normal	800uA
32nm LVT	5.9nA

2. 6-bit to 64-bit Decoder

Another application we considered was the 6-bit Decoder. For body biasing purposes the body terminal of each NMOS and PMOS transistors are tied to the V_{bodyn} and V_{bodyp} terminals respectively. as shown for the AND gate in Fig 4.

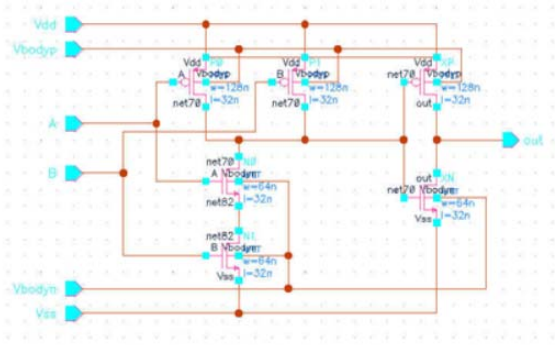


Fig 4. AND gate with body bias terminarls

The results after using body bias voltage are presented in Fig 5. The simulation shows that the effect of body bias is not that great compared to other methods such as power gating or dual threshold voltage techniques. Although, according to the reference search,[2][3] body bias may significantly decrease the leakage current, form the observations we have made, body biasing at best should be used as a supplementary method to other methods such as input vector forcing or power gating.

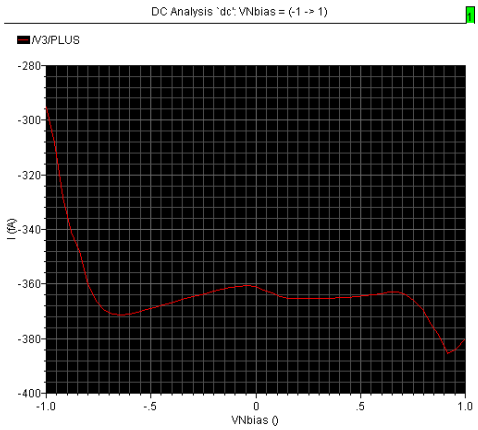


Fig 5. Leakage current with body biasing

3. Clock Tree

A clock tree was selected to simulate. The simulation included the sleep transistor size and delay and leakage current. Also I did the simulation of the leakage current versus the body bias.

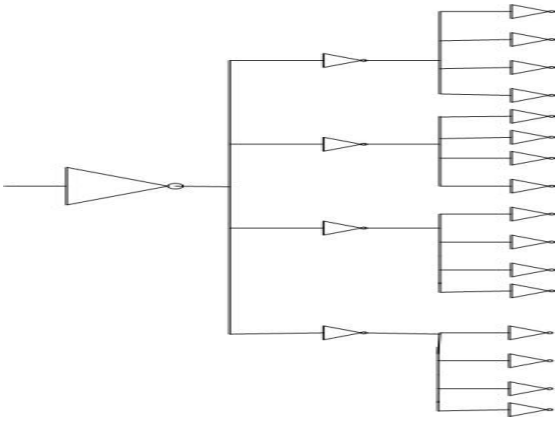
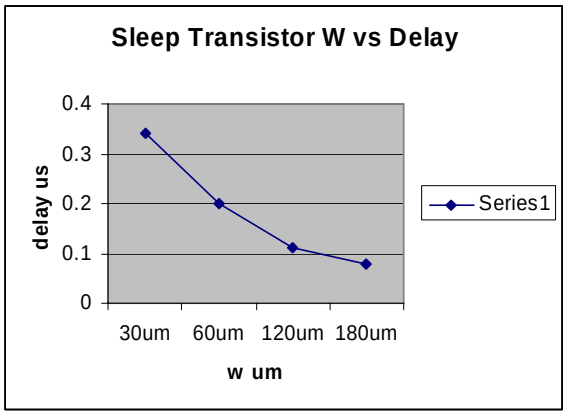


Fig 6. A Clock Tree

The clock is entirely made up of invertors, because leakage the transistor stacks, it more susceptible to the sub-threshold current.



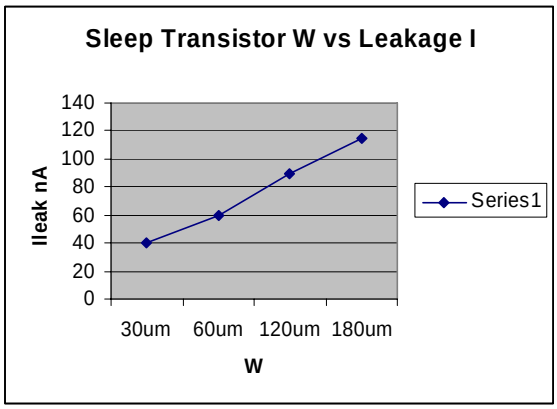
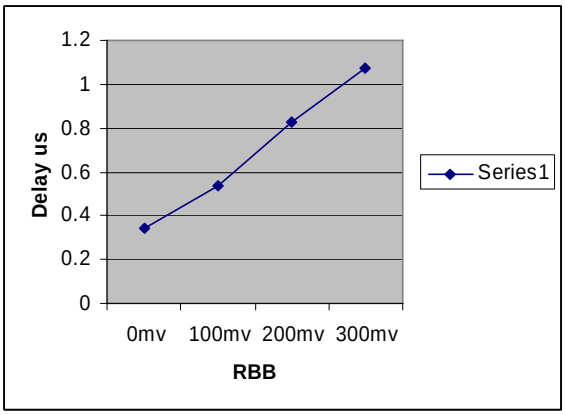


Fig 7. Sizing of the sleep transistor



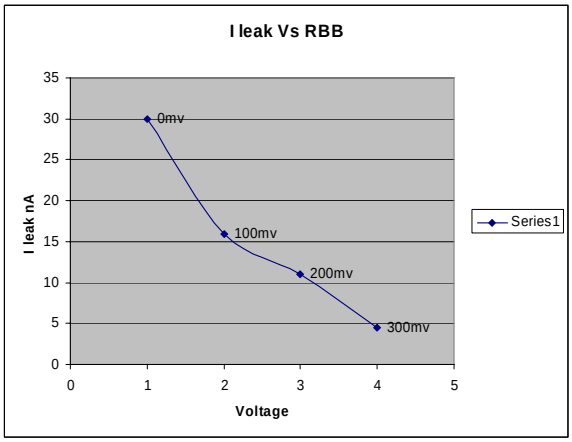


Fig. 8. Simulation of reverse body bias.

Increasing the sleep transistor width can cause a larger leakage current, but it reduces the delay at the same time. There is a trade off between delay and leakage power.

Increasing the Reverse body bias, it will worse the delay but reduce the leakage current.

4. 32 Bit Carry Look-ahead Adder

This 32 bits adder uses the dynamic domino logic and implemented the dual V_t technology.

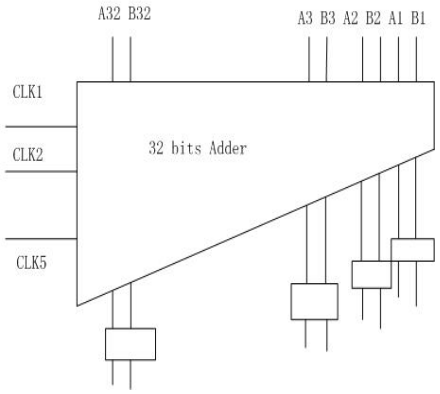


Fig 9. 32bits adder

	Eva.	Pre.	Leakage input=1	Leakage Input =0
HVT	40ps	29ps	32fA	1.2uA
LVT	24ps	17ps	300fA	53uA
DVT	22ps	28ps	270fA	39uA

Table 2: The result of adder with L=45nm

	Eva.	Pre.	Leakage input=1	Leakage Input =0
HVT	32ps	26ps	45fA	1.8uA
LVT	18ps	13ps	430fA	66uA
DVT	17ps	26ps	360fA	54uA

Table 3: The result of adder with L =32nm

From the table we can see that the Dual V_t adder have a closer evaluation delay as the low V_t adder, and a closer precharge delay as the High V_t Adder.

For the worst case leakage current, Dual VT device is between High V_t and Low V_t. Actually the leakage current greatly related to the size of the high V_t device. I didn't do the optimal design to fine the largest size of High V_t device under a certain tolerance of delay. Sizing for this domino logic is simply optimize base the same stage logic block, which is much easier than doing that in power gating where people need to consider the whole block.

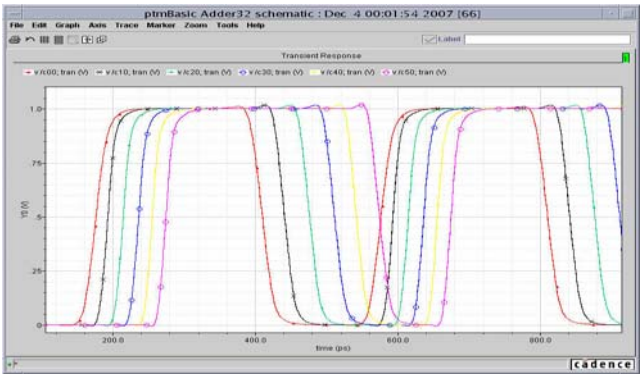
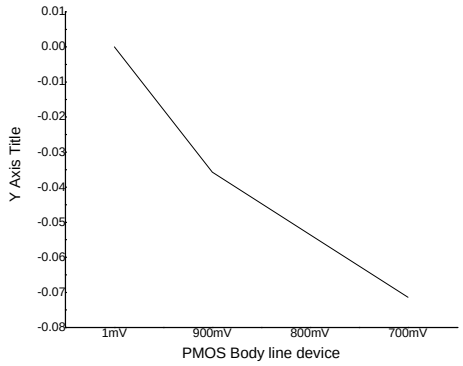


Fig 10, Clock diagram

The clock diagram can see that the uneven of the charging and discharging cause the clock width expending. If the falling edge and the rising edge overlap, that would a logic error. In order to enhance the precharge speed , I considered use a forward body bias on the high Vt device.



Fig(8) Delay vs Forward body bias

The precharge delay only improves a little bit fewer than 10%. I think that is because the adder pull down network have a transistor stack, so the body bias effect is not significant compare with the clock tree. The clock tree is simply made of inverter without a transistor stack.

IV. Conclusion

The effect of various methods of leakage reduction on several commonly used applications was examined. The results indicate that the power gating is the most effective in reducing leakage. However it has the problem of inducing delays in to the system. There form it is recommended that a combination of power gating and dual threshold should be used for large scale applications.

References

[1] J. Kao, A. Chandrakasan. Dual-Threshold Voltage Techniques for Low-Power Digital Circuits, IEEE Journal of Solid-state Circuits, Vol. 35, No. 7, July 2000.

[2] C. Neau, K. Roy. Optimal body bias selection for leakage improvement and process compensation over different technology generations, Low Power Electronics and Design, 2003. ISLPED '03. Proceedings of the 2003 International Symposium on 25-27 Aug. 2003 Page(s):116 - 121

[3] Zhengya Zhang, Zheng Guo, Active Leakage control with Sleep Transistors and Body Bias, Berkeley University, 2004.